

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/498,556

DATE: 06/12/2001
TIME: 15:06:48

Input Set : N:\Crf3\RULE60\09498556.txt
Output Set: N:\CRF3\06122001\I498556.raw

ENTERED

SEQUENCE LISTING

5 (1) GENERAL INFORMATION:

8 (i) APPLICANT: Vlasuk, George Phillip
9 Stanssens, Patrick Eric Hugo
10 Messens, Doris Hilda Lieven
11 Lauwerss, Marc Josef
12 Laroche, Yves Rene
13 Jemperis, Laurent Stephane
14 Gensemans, Yannick Georges Jozef
15 Hoyle, Matthew
16 Bergum, Peter W.

17 (ii) TITLE OF INVENTION: NEMATODE-EXTRACTED SERINE PROTEASE
18 INHIBITORS AND ANTICOAGULANT
21 PROTEIN

24 (iii) NUMBER OF SEQUENCES: 156

27 (iv) CORRESPONDENCE ADDRESS:

29 (A) ADDRESSEE: Lyon & Lyon
30 (B) STREET: 634 West Fifth Street
31 Suite 400
32 (C) CITY: Los Angeles
33 (D) STATE: California
34 (E) COUNTRY: U.S.A.
35 (F) ZIP: 90011

38 (v) COMPUTER READABLE FORM:

40 (A) MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
41 storage
42 (B) COMPUTER: IBM Compatible
43 (C) OPERATING SYSTEM: IBM P.C. DOS 5.0
44 (D) SOFTWARE: Word Perfect 5.1

47 (vi) CURRENT APPLICATION DATA:

49 (A) APPLICATION NUMBER: US/09/498,556
50 (B) FILING DATE: 04-Feb-2000

51 (vii) PRIOR APPLICATION DATA:

55 (A) APPLICATION NUMBER: 08-804,456
56 (B) FILING DATE: April 17, 1997
58 (A) APPLICATION NUMBER: PCT/US95/13231
61 (B) FILING DATE: October 1, 1995
63 (A) APPLICATION NUMBER: 08-456,399
64 (B) FILING DATE: June 5, 1995
66 (A) APPLICATION NUMBER: 08-456,397
67 (B) FILING DATE: June 5, 1995
70 (A) APPLICATION NUMBER: 08-460,380
71 (B) FILING DATE: June 5, 1995
73 (A) APPLICATION NUMBER: 08-461,965
74 (B) FILING DATE: June 5, 1995
76 (A) APPLICATION NUMBER: 08/326,110
77 (B) FILING DATE: October 18, 1994

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/498,556

DATE: 06/12/2001
 TIME: 15:06:48

Input Set : N:\Crf3\RULE60\09498556.txt
 Output Set : N:\CRF3\06122001\I498556.raw

| | | |
|-----|---|--|
| 80 | (viii) ATTORNEY/AGENT INFORMATION: | |
| 81 | (A) NAME: BIGGS, CYNTHIA L. | |
| 82 | (B) REGISTRATION NUMBER: 30,158 | |
| 83 | (C) REFERENCE DOCKET NUMBER: 216/270 | |
| 84 | | |
| 85 | (ix) TELECOMMUNICATION INFORMATION: | |
| 86 | (A) TELEPHONE: (213) 489-1600 | |
| 87 | (B) TELEFAX: (213) 455-0440 | |
| 88 | (C) FAX: 62-5410 | |
| 89 | | |
| 90 | (2) INFORMATION FOR SEQ ID NO: 1: | |
| 91 | (i) SEQUENCE CHARACTERISTICS: | |
| 92 | (A) LENGTH: 111 base pairs | |
| 93 | (B) TYPE: nucleic acid | |
| 94 | (C) STRANDEDNESS: single | |
| 95 | (D) TOPOLOGY: linear | |
| 96 | | |
| 97 | (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1: | |
| 98 | AAGGCATAAC GAAAGTGGCG TGAATGAA TGGCTTGACG ACTGTGGAAC TCAGAACCCA 60 | |
| 99 | TGCGAGGCCA AATTTATGA GAAATGCGT GAGGAGGAAG ATCCGATATG CGGCTCACGT 120 | |
| 100 | GGTTGTTATP TATTTCTGT TTGCTATG AAAGACGGAT TCTACAGAGA CACGGTGATC 180 | |
| 101 | GGCGACTGTC TAAAGGAGA AGATTCGAC CAACATGAGA TTATACATGT CTGA 234 | |
| 102 | | |
| 103 | (2) INFORMATION FOR SEQ ID NO: 2: | |
| 104 | (i) SEQUENCE CHARACTERISTICS: | |
| 105 | (A) LENGTH: 124 base pairs | |
| 106 | (B) TYPE: nucleic acid | |
| 107 | (C) STRANDEDNESS: single | |
| 108 | (D) TOPOLOGY: linear | |
| 109 | | |
| 110 | (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2: | |
| 111 | AAGGCATAAC GAAAGTGGCG TGAATGAA TGGCTTGACG TCTGPGAAC TAAGAAGGCA 60 | |
| 112 | TGCGAGGCCA AATTTATGA GAAAGGAGG GAAGATCCGA TATGCGGATC ATTTCTGT 120 | |
| 113 | CCGGGTCGGG GATTTGGT ATGAGAGAG GGATTCTACA GAGACACGGT GATCGGCGAC 180 | |
| 114 | TGTGTTAAGT AAGAAGATG GAGGAAAT GAGATTATAC ATGTCTGA 228 | |
| 115 | | |
| 116 | (2) INFORMATION FOR SEQ ID NO: 3: | |
| 117 | (i) SEQUENCE CHARACTERISTICS: | |
| 118 | (A) LENGTH: 161 base pairs | |
| 119 | (B) TYPE: nucleic acid | |
| 120 | (C) STRANDEDNESS: single | |
| 121 | (D) TOPOLOGY: linear | |
| 122 | | |
| 123 | (vi) ORIGIN OF SOURCE: | |
| 124 | (A) ORGANISM: Amyloplasma caninum | |
| 125 | | |
| 126 | (ix) FEATURE: | |
| 127 | (A) NAME/KEY: Coding Sequence | |
| 128 | (B) LOCATION: 32...371 | |
| 129 | | |
| 130 | (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3: | |
| 131 | GAATTCCGCT AATCTAAC A ATG AAG ATG CTT TAC GCT ATC GCT ATA ATG 51 | |
| 132 | Met Lys Met Leu Tyr Ala Ile Ala Ile Met | |
| 133 | 1 5 10 | |
| 134 | TTT CTC CTG GTA TAA TTA TGC A AC GCA AGA ACA GTG AGG AAG GCA TAC 99 | |
| 135 | Phe Leu Leu Val Ser Leu Cys Ser Ala Arg Thr Val Arg Lys Ala Tyr | |
| 136 | 11 20 25 | |
| 137 | CCG GAG TGT GGT GAS AAT GAA TGG CTC GAC GAC TGT GGA ACT CAG AAG 147 | |

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/498,556

DATE: 06/12/2001
TIME: 15:06:48

Input Set : N:\Crf3\RULE60\09498556.txt
Output Set: N:\CRF3\06122001\I498556.raw

164 Pro Glu Cys Gly Gln Asn Glu Trp Leu Asp Asp Cys Gly Thr Gln Lys
165 30 35 40
167 CCA TGC GAG GC¹ AAG TGC AAT GAG GAA CCT CCT GAG GAG GAA GAT CCG 195
168 Pro Cys Glu Ala Lys Cys Asn Gln Glu Pro Pro Glu Glu Asp Pro
169 45 50 55
171 ATA TGC CGC TCA CCT GGT TGT TTA PTA CCT GCT GCT TGC GTA TGC AAA 243
172 Ile Cys Arg Ser Arg Cys Ser Leu Pro Ala Cys Val Cys Lys
173 60 65 70
175 GAC GGA TTC TAG AGA GAC ACG GTG ATC GGC GAC TGT GTT AGG GAA GAA 291
176 Asp Gly Phe Tyr Arg Asp Thr Val Ile Gly Asp Cys Val Arg Gln Glu
177 75 80 90
179 GAA TGC GAC GAA CAT AGG ATT ATA CAT GTC T GAAAGAGAAA GCAACAATAA CC 344
180 Glu Cys Asp Gln His Gln Ile Ile His Val
181 90 100
183 AAAGGGTTCCA ATTCTGGTC TGTAATACTG CTAGTIGGAT GTCTCTTTG CGTCCGAATA 404
185 GTTTTAGTCG ATTTAAAGTA AGAAATTCTG CTGGAGAGAA TAAAGCTTTC CAACTCC 461
187 (2) INFORMATION FOR SEQ ID NO: 4:
188 (i) SEQUENCE CHARACTERISTICS:
189 (A) LENGTH: 77 amino acids
190 (B) TYPE: amino acid
191 (C) TOPOLOGY: linear
192 (ii) MOLECULE TYPE: peptide
193 (vi) ORIGINAL SOURCE:
201 (A) ORGANISM: Ancylostoma caninum
203 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:
204 Lys Ala Tyr Phe Gln Cys Gly Glu Asn Gln Trp Leu Asp Asp
205 1 5 10 15
206 Cys Gly Thr Gln Lys Pro Cys Glu Ala Lys Cys Asn Gln Glu
207 15 20 25
211 Pro Pro Gln Gln Glu Asp Pro Ile Cys Arg Ser Arg Gly Cys
212 30 35 40
214 Leu Leu Pro Pro Ala Cys Val Cys Lys Asp Gly Phe Tyr Arg
215 45 50 55
217 Asp Thr Val Ile Gln Asp Cys Val Arg Gln Glu Glu Cys Asp
218 60 65 70
220 Gln His Gln Ile Ile His Val
221 75
224 (2) INFORMATION FOR SEQ ID NO: 5:
226 (i) SEQUENCE CHARACTERISTICS:
227 (A) LENGTH: 455 base pairs
228 (B) TYPE: nucleic acid
229 (C) STRANDEDNESS: single
230 (D) TOPOLOGY: linear
233 (vi) ORIGINAL SOURCE:
235 (A) ORGANISM: Ancylostoma caninum
237 (ix) FEATURE:
239 (A) NAME KEY: Coding Sequence
240 (B) LOCATION: 22...315
242 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/498,556

DATE: 06/12/2001

TIME: 10:06:48

Input Set : N:\Crf3\RULE60\09498556.txt

Output Set: N:\CRF3\06122001\I498556.raw

| | | |
|-----|--|-----|
| 244 | GAATTCGCT ACTACTCAAC A ATG AAG ATG CTT TAC GCT ATC GCT ATA AAG | 51 |
| 245 | Met Lys Met Leu Tyr Ala Ile Ala Ile Met | |
| 246 | 1 6 10 | |
| 248 | TTT CTC CTG GTG TCA TTA TGC AGC ACA AGA AGA GTG AGG AAG GCA TAA | 99 |
| 249 | Phe Leu Leu Val Ser Leu Cys Ser Thr Arg Thr Val Arg Lys Ala Tyr | |
| 250 | 15 20 25 | |
| 252 | CCG GAG TGT GGT GAG AAT GAA TGG CTC GAC GTC TGT GCA ACT AAG ARG | 147 |
| 253 | Pro Glu Cys Gly Glu Asn Glu Trp Leu Asp Val Cys Gly Thr Lys Lys | |
| 254 | 30 35 40 | |
| 256 | CCA TGC GAG GTC AAG TGC AGT GAG GAA GAG GAA GAT CGG ATA TGC | 195 |
| 257 | Pro Cys Gli Ala Lys Cys Ser Glu Glu Glu Glu Asp Pro Ile Cys | |
| 258 | 45 50 55 | |
| 262 | CGA TCA TTP TCT TGT CGG GGT CCT GCT GCT TGC GTA TGC GAA GAC GAA | 243 |
| 263 | Arg Ser Phe Ser Cys Pro Gly Pro Ala Ala Cys Val Cys Glu Asp Gly | |
| 264 | 60 65 70 | |
| 266 | TTC TAC AGA GAC ACG GTG ATC GGC GAC TAA GTT ARG GAA GAA GAA TGC | 291 |
| 267 | Phe Tyr Arg Asp Thr Val Ile Gly Asp Cys Val Lys Glu Glu Glu Cys | |
| 268 | 75 80 85 90 | |
| 270 | GAC CAA CAT GAG ATT ATT CAT GTC TGAAGGAGAG AGCAGTAATA ACCAAAGGTT C | 346 |
| 271 | Asp Gln His Glu Ile Ile His Val | |
| 272 | 95 | |
| 274 | CAACTTCCG TTTACAAAAT CGCTAGTTGG APTTTCCTT TCGCTCGAA TAGTTTTAGT | 406 |
| 275 | TGATATTAAAG TAIIAACCTCG TGTTGAAGAG AATAALAGCTT TCGAACTT | 455 |
| 279 | (2) INFORMATION FOR SEQ ID NO: 6: | |
| 281 | (i) SEQUENCE CHARACTERISTICS: | |
| 283 | (A) LENGTH: 75 amino acids | |
| 284 | (B) TYPE: amino acid | |
| 285 | (C) TOPOLOGY: linear | |
| 287 | (ii) MOLECULE TYPE: peptide | |
| 289 | (vi) ORIGINAL SOURCE: | |
| 291 | (A) ORGANISM: Ancylostoma caninum | |
| 293 | (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6: | |
| 295 | Lys Ala Tyr Pro Glu Cys Gly Glu Asn Glu Trp Leu Asp Val Cys Gly | |
| 296 | 1 6 10 14 | |
| 298 | Thr Lys Lys Pro Cys Glu Ala Lys Cys Ser Glu Glu Glu Glu Asp | |
| 299 | 20 25 30 | |
| 301 | Pro Ile Cys Arg Ser Phe Ser Cys Pro Gly Pro Ala Ala Cys Val Cys | |
| 302 | 35 40 45 | |
| 304 | Glu Asp Gly Phe Tyr Arg Asp Thr Val Ile Gly Asp Cys Val Lys Glu | |
| 305 | 50 55 60 | |
| 307 | Glu Glu Cys Asp Gln His Glu Ile Ile His Val | |
| 308 | 65 70 75 | |
| 310 | (2) INFORMATION FOR SEQ ID NO: 7: | |
| 312 | (i) SEQUENCE CHARACTERISTICS: | |
| 314 | (A) LENGTH: 61 amino acids | |
| 315 | (B) TYPE: amino acid | |
| 316 | (D) TOPOLOGY: linear | |
| 318 | (ii) MOLECULE TYPE: peptide | |
| 320 | (vi) ORIGINAL SOURCE: | |

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/498,556

DATE: 06/12/2001

TIME: 15:06:48

Input Set : N:\Crf3\RULE60\09498556.txt

Output Set: N:\CRF3\06122001\I498556.raw

321 (A) ORGANISM: Ancylostoma caninum
 324 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:
 326 Arg Thr Val Arg Lys Ala Tyr Pro Glu Cys Gly Glu Asn Glu Trp Leu
 327 1 . . . 10 15
 328 Asp Asp Cys Gly Thr Gln Lys Pro Cys Glu Ala Lys Cys Asn Glu Glu
 330 20 25 30
 331 Pro Pro Glu Ser Glu Asp Pro Ile Cys Arg Ser Arg Gly Cys Leu Leu
 333 30 40 45
 334 Pro Pro Ala Cys Val Cys Lys Asp Gly Phe Tyr Arg Asp Thr Val Ile
 336 40 50 60
 337 Gly Asp Cys Val Arg Glu Glu Cys Asp Gln His Glu Ile Ile His
 338 55 65 75 80
 341 Val
 346 (2) INFORMATION FOR SEQ ID NO: 7:
 348 (i) SEQUENCE CHARACTERISTICS:
 350 (A) LENGTH: 79 amino acids
 351 (B) TYPE: amino acid
 352 (C) TOPOLOGY: linear
 354 (ii) MOLECULE TYPE: peptide
 356 (vi) ORIGINAL SOURCE:
 358 (A) ORGANISM: Ancylostoma caninum
 360 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:
 362 Arg Thr Val Arg Lys Ala Tyr Pro Glu Cys Gly Glu Asn Glu Trp Leu
 363 1 . . . 10 15
 364 Asp Val Cys Gly Thr Lys Lys Pro Cys Glu Ala Lys Cys Ser Glu Glu
 366 20 25 30
 368 Glu Glu Glu Asp Pro Ile Cys Arg Ser Phe Ser Cys Pro Gly Pro Ala
 369 30 40 45
 371 Ala Cys Val Cys Glu Asp Gly Phe Tyr Arg Asp Thr Val Ile Gly Asp
 372 40 50 60
 374 Cys Val Lys Glu Glu Glu Cys Asp Gln His Glu Ile Ile His Val
 375 55 65 75
 380 (2) INFORMATION FOR SEQ ID NO: 8:
 382 (i) SEQUENCE CHARACTERISTICS:
 384 (A) LENGTH: 71 base pairs
 385 (B) TYPE: nucleic acid
 386 (C) STRANDEDNESS: single
 387 (D) TOPOLOGY: linear
 389 (vi) ORIGINAL SOURCE:
 391 (A) ORGANISM: Ancylostoma ceylanicum
 393 (ix) FEATURE:
 395 (A) NAME KEY: Coding Sequence
 396 (B) LOCATION: 21...590
 399 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:
 401 GAATTCACTA TTATCCAACA ATG GCG GTG CTT TAT TCA GTA GCA ATA GCG 50
 402 Met Ala Val Leu Tyr Ser Val Ala Ile Ala
 403 1 . . . 5 10
 404 TTA CTA CTG GTA TAA CAA TAC AGT GGG AAA CCG AAC AAT GTG ATG ACT 98
 405 Leu Leu Val Ser Gln Cys Ser Gly Lys Pro Asn Asn Val Met Thr

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/498,556

DATE: 06/12/2001

TIME: 15:06:49

Input Set : N:\Crf3\RULE60\09498556.txt

Output Set: N:\CRF3\06122001\I498556.raw

L:43 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]
L:50 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]
L:2909 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:66
L:3323 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:67
L:3353 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:68
L:3155 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:78
L:4185 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:79
L:2110 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:80
L:3335 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:81
L:4160 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:82
L:2283 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:83
L:3712 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:84
L:3336 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:85
L:3362 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:86
L:3363 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:87
L:3351 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:118
L:3370 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:119
L:4163 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:129
L:4189 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:130
L:4165 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:131
L:4231 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:132
L:4254 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:133
L:4176 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:134
L:4177 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:134
L:4301 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:135
L:4104 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:136
L:4326 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:137
L:4327 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:137
L:4328 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:138
L:4402 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:139
L:4438 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:140
L:4446 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:141
L:4471 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:142
L:4484 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:143
L:4417 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:144
L:4541 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:145
L:4564 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:146
L:4587 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:147
L:4611 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:148
L:4634 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:149
L:4657 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:150
L:4686 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:151
L:4760 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:152
L:4716 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:153
L:4710 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:154

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/498,556

DATE: 06/12/2001

TIME: 15:06:49

Input Set : N:\Crf3\RULE60\09498556.txt
Output Set: N:\CRF3\06122001\I498556.raw

L:4773 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:155

L:4797 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:156

L:4819 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:157

L:4841 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:158